Negative Performance Information Causes Asymmetrical Evaluations and Elicits Strong Responsibility Attributions

Olsen, Asmus Leth

Publication date:
2015

Negative Performance Information Causes Asymmetrical Evaluations and Elicits Strong Responsibility Attributions

Asmus Leth Olsen
Department of Political Science
University of Copenhagen
www.asmusolsen.com
ajlo@ifs.ku.dk

Abstract

The negativity bias challenges our understanding of how both citizens and managers respond to performance information. It posits that negative information has a stronger impact than positive information of the same magnitude. However, studies of bias in public administration have revealed mixed results. We conduct the first set of experiments with the sole purpose of estimating the negativity bias in citizens' responses to performance information. Moreover, we extend the analysis beyond evaluations and outline a theory and test of how negative performance information also elicits stronger responsibility attributions among citizens: a search for the underlying causes for poor performance. We investigate the negativity bias with three experiments embedded in two large nationally-representative samples of citizens (n=1,559 and n=1,013). Within a conservative equivalence framing framework, we find very strong evidence of a large negativity bias. A direct replication confirms this conclusion. In addition, we find that citizens are much more likely to spontaneously engage in attributions about responsibility after being exposed to negative performance data. The findings have implications for our understanding of both citizens and managers responses to performance information and provides a strong empirical connection between negativity bias and blame avoidance.

Keywords: negativity bias; responsibility attribution; performance information; survey experiments; blame avoidance

Acknowledgment:
Ideas expressed in the manuscript has benefited from helpful comments and suggestions from Oliver James, Sebastian Jilke, Martin Vinæs Larsen, and Gregg van Ryzin. Parts of the results were presented at a pre-conference on experimental public administration at the Public Management Research Conference (PMRC) June 11-13, 2015, Minneapolis, Minnesota. The author would like to thanks participants for helpful comments and suggestions.

Funding:
The research was funded by two grants: 1) Grant no. 0602-02577B from the Danish Council for Independent Research and 2) a grant from the Master of Public Governance (MPG) at the University of Copenhagen and Copenhagen Business School in Denmark. Project title: "The Psychology of the Public Sector".

Introduction

The negativity bias has recently gained considerable attention in the field of public administration, particularly regarding the study of attitudinal and behavioral responses to performance information (James and John 2007; Boyne et al. 2009; James 2011a; 2011b; James and Moseley 2015; Olsen 2015). The negativity bias posits a fundamental asymmetrical response to positive and negative information in which "negative events are more salient, potent, dominant in combinations, and generally efficacious than positive events" (Rozin and Royzman 2001, 297). The implications and importance of understanding a negativity bias go beyond the context of how citizens draw on performance information. Most notably, negativity bias is seen as the driving micro-foundation for blame avoidance among administrative and political leaders (Weaver 1986: 373; Marvel 2014; Moynihan 2012; Hood 2007; 2010; Dixon and Hood 2010; Charbonneau and Bellavance 2012; Nielsen and Bækgaard 2015). In short, if the effect of negative and positive information is highly skewed towards the former, then performance information will systematically increase the assignment of blame while having little or no effect on the assignment of praise and credit for good performance. The down-stream effect in the political-administrative system is that managers and politicians will primarily follow a mini-max strategy and be more concerned with avoiding bad performances than with striving for excellence (Hood and Dixon 2010). However, while psychology finds consistent support for a negativity bias in human perception, emotional responses, decision making, behavior, and memory (Baumeister et al. 2001; Rozin and Royzman 2001), the findings in public administration have been more mixed. James (2011b, 414) argues that, "More investigation of possible difference in magnitude of effect between information about good and bad performance is merited." Furthermore, we have no direct evidence of how a negativity bias would affect citizens' attribution of responsibility for the observed performance.

In this article, we conduct three experiments in two large and nationally-representative samples (n=1,559 and n=1,013) which directly (a) address the question of a negativity bias in citizens responses to performance information and (b) provide the first test of how negative performance elicits responsibility attribution. For the study of the negativity bias in evaluations we draw on a modified equivalence framing experiment in which subjects are provided substantively equivalent descriptions of the performance level but with varying valence, i.e. negative, neutral, and positive (Levin, Schneider, and Gaeth 1998). It solves three fundamental problems of the current research: First, it allows us to compare the effect of "good" and "poor" performance that only differs in valence, but is qualitatively the same. Second, it provides a neutral level of performance which allows us to estimate the asymmetrical effect of positive and negative information relative to a neutral benchmark. Thirdly, it is a hard test of the negativity bias which implies that any confirming evidence will be more credible. For the study of how negative performance information elicits attributions we draw on a long tradition in social psychology of studying causal triggers (Roese 1997, Wiener 1985). Social psychology posits that negative feedback is fundamental to the search for explanations: Why did something go wrong? We use a novel method of open-text response to track and code the spontaneous attributions which citizens make after being exposed to performance information of varied valence.

We find consistent and strong evidence of a negativity bias in citizens' responses to various performance information. The asymmetry is striking: Citizens responses to positive performance information is very similar to their responses to neutral information - only if performance information is negative do we observe strong and robust differences. The pattern is found across two nationally-representative surveys as we have also conducted an independent replication. Furthermore, we find that the negativity bias extends far beyond an immediate asymmetry in citizens evaluations. Most strikingly, we find that citizens

spontaneously produce causal attributions if exposed to negative valence performance information. In short, they are far more likely to generate explanations of why organizations obtain negative results than if results are positive. The findings have a number of implications for future research: First, the findings indicate that the negativity bias is a basic human feature which is likely to play a role for perceptions and decision-making institutional features which amplify negative information (Hood 2007; 2010, 13). Second, it lends credibility to recent studies on blame avoidance among managers and politicians: a strong negativity bias increases the stakes at hand for avoiding negative performances. It also raises important questions about politicians' strategies to avoid blame (Hood and Dixon 2010). Third, the findings forces us to investigate if the negativity bias is in fact a "bias" in a normative sense – maybe citizens are better off by narrowly focusing on negative feedback? These implications are outlined further in the final section.

The Negativity Bias and Public Performance Information

The negativity bias posits that negative events have a greater impact on the individual than positive events of the same magnitude. Psychology finds consistent support for a negativity bias in human perception, emotional response, decision making, behavior, and memory (Baumeister et al. 2001; Rozin and Royzman 2001). Evolutionary explanations emphasize the potential adaptive value of responding in an asymmetrical fashion to negative events. Throughout human history negative events have been more extreme, irreversible, contagious, and potentially life-threatening (Hibbing, Smith, and Alford 2014). A negativity bias could therefore be seen as ecologically rational because, with the limits of bounded rationality in mind (Simon 1955), responding more strongly to negative inputs is a rule of thumb that serves us well in most instances.

In public administration the negativity bias has largely been seen as problematic and unproductive because it skews citizens' evaluation of performance information which, in turn, affects how credit and blame is assigned to politicians and managers (Boyne et al. 2007; Hood 2007; 2010). The idea of a negativity bias in citizens' evaluations is at odds with a symmetrical view of the impact caused by performance information. Clearly, if the negativity bias is important for the effect of performance information, then citizens' attention, attitudes, and decision-making will mainly be affected by negative performance information, which can have important downstream effects on politicians and managers behavior. However, while psychology provides a strong empirical and theoretical foundation for the bias, we cannot directly draw implications from these findings to the field of public administration and, more specifically, the study of performance information. For instance, in psychology most studies of negativity bias focus on how the individual weighs negative and positive events which they personally experience or which they observe occurring for other individuals (Fiske 1980). For the case of performance information, we primarily study how citizens draw inferences from information about "abstract" entities such as organizations, governments, or countries. Studies of the negativity bias are therefore merited in a political-administrative context.

If we take a closer look on studies of the negativity bias in public administration we find a number of both observational and experimental studies with mixed findings. Some of these issues stem from the fact that, in most of these studies, the evidence for or against a negativety bias is often indirect and not the primary target of the study. James and John (2007) relies on data from the Comprehensive Performance Assessments (CPA) of local authorities in England. By connecting local CPA scores with the vote share of the incumbent they find a "*clear negativity bias, with the category of poor performance resulting in punishment but excellent performance not being equally rewarded*" (James and John 2007,

577). Along the same lines, Boyne et al. (2009, 1282) also analyze CPA scores, incumbent votes, and citizen satisfaction surveys and conclude that "*Negativity bias exists. Only the difference between low performance and anything better matters*". James (2011a) combines observational and experimental evidence to test differences in how expectations form citizens' evaluations of performance information, concluding that "*negativity bias is evident with information about poor performance having a larger effect*". One potential criticism of these studies is that the negative and positive information being compared is in fact qualitatively different - for instance if citizens do not understand "poor" and "excellent" performance as equivalent labels with opposing valence. Psychologists have long struggled with separating magnitude or "extremity" of information from its valence as negative information is often more extreme, unexpected, and infrequently observed (Rozin and Royzman 2001). A related issue is to define the boundaries we draw between negative and positive information. Often citizens use their expectations as a reference point as to why they "often take good government for granted and pay more attention to insufficient performance" (Yang and Holzer 2006, 116). The observational studies also highlight these weak points and note how "negativity bias in political behavior would also seem to merit further research at the individual level" (John and James 2007, 578).

While the observational studies generally find a negativity bias, the experimental evidence is more mixed. James (2011b) experimentally assigned positive and negative information in both a lab and a field setting and "did not find clear evidence of negativity bias", although the study notes that the null finding may be explained by a divergence between the experimental treatments and the actual performance of the units citizens were evaluating. James and Moseley (2014) set out to compare absolute and relative performances, which also indirectly allows them to evaluate difference in how positive and negative information affect citizens. They find some support for a negativity bias insofar as

absolute bad performance decreases citizens' satisfaction while good performance has no effect. However, for relative performance the findings are inconclusive and they note how "*more research on negativity bias within a single integrated experiment would help confirm these results…*" (James and Moseley 2015, 13). Finally, Olsen (2015) relies on an equivalence framing experiment to study how citizens evaluate public hospitals given information about either a certain level of "satisfaction" vs. a level of "dissatisfaction". The study finds that citizens respond much more negatively to a "10% dissatisfaction rate" than a logically equivalent "90% satisfaction rate". However, showing that citizens respond differently does necessarily imply a negativity bias, but might as well be due to a positivity bias. We need a neutral point for comparison in order to find out if the "negative" or the "positive" information is causing the asymmetry. We cannot determine which of the two is driving the difference without a benchmark to compare the relative effect of positive and negative.

Here we will adopt a *modified* equivalence framing setup to solve the criticisms and mixed findings of the existing studies. Equivalence frames are defined as "objectively equivalent descriptions of the same problem" (Levin, Schneider, and Gaeth 1998, 150). For instance, many performance indicators can be labeled with either a positive or negative valence label: death rates/survival rates, satisfaction/dissatisfaction, pass/fail, employment/unemployment etc. While holding constant the numerical level of performance we can switch labels to change the information's valence. For instance, we can present some citizens with a "10% fail rate" for a public school while others are exposed to a "90% pass rate". Logically, the two pieces of information are equivalent, yet their difference in valence is likely to provide very different evaluations of the school to those exposed to the information. The underlying theoretical mechanism rests on individual encoding of information given its descriptive valence (March and Simon 1958; Levin and Gaeth 1988;

Quattrone and Tversky 1988). Encoding affects the associative memory by increasing the accessibility of valence-consistent information, i.e., positive valence performance information elicits positive associations while negative valence performance information elicits negative associations. For instance, when we told about a *fail* we are immediately reminded about everything that can be wrong with a public school, such as lazy students, incompetent teachers, corrupt management, or rundown facilities due to a lack of funds. Positive and negative associations cause a "valence-consistent shift" in an individual's evaluation (Levin, Schneider, and Gaeth 1998). Accordingly, even when the underlying performance is identical, the valence of the information will affect evaluations: Negative-valence information obtains a more negative evaluation than positive information with the same substantive information, e.g. a "10% *fail* rate" for a public school vs. a "90% *pass* rate". The equivalence framing perspective provides solutions for the problems identified in the existing studies. First, by presenting citizens with logically-equivalent numerical information and only switching the valence of the information's label we overcome concerns about the positive and negative information being qualitatively different on other dimensions. Second, we propose a modified equivalence frame where we compare the effect of negative and positive information relative to a neutral midpoint where citizens are informed by both the negative and the positive information. This allows us to identify if the negative frame is, in fact, causing the asymmetry. Finally, the equivalence framework provides a hard and conservative theoretical underpinning for testing the negativity bias. Theoretically, it requires that we identify an asymmetry between positive and negative performance information even when the underlying information is identical. It implies that if we find evidence in favor of a negativity bias within this framework, it will lend strong credibility to the strength and relevance of the asymmetrical effect of negative information on citizens' evaluations. From this basis we can now formulate our first main hypothesis:

Negativity bias in evaluations: *Negative valence performance information will have a greater impact on citizens' evaluations of a public service than positive valence performance information of the same magnitude relative to neutral performance information.*

Beyond Evaluations: Negative Information Elicits Responsibility Attributions

The asymmetrical effect of positive and negative information reaches well beyond the dimension of evaluation and judgment. Reviews of the negativity bias stress how negative input dominates or exerts widely different responses than positive information on a wide range of dimensions. These include perception, attention, physiological arousal, memory, impression formation and generally, responses to negative stimuli, are often more differentiated and elaborate than to positive ones (Baumeister et al. 2001, Rosin and Rozyman 2001). From a political-administrative perspective the most important one of these is the difference in how negative and positive feedback prompts an attributional process (Kelly 1973; Wiener 1984; Roese 1997). Greater information processing is likely to produce more elaborate, thorough, and extensive cognitive interpretations, which includes attributions about the information (Baumeister et al. 2001, 340). In other words, negative input elicits more attributional activity (Rozin and Royzman 2001: 301): We think about why something went wrong, not why it went well. This basic idea is far from new to public administration. At an organizational level there has long been a focus on how negative feedback plays a distinct role beyond evaluations. We broadly encounter the view that performance is coded as either a loss or gain, success or failure, negative or positive, and good or bad - this, in turn, affects how managers and organizations respond (March and Simon 1958; Meier et al. 2015). Attributions about performance are initiated as soon as dissatisfaction and failure is observed. It becomes a "*stimulus to search for lessons*" (Rose 1991), a trigger for "*problemistic search*", or initiates "trial-*and-error experimentation*" (Cyert & March 1963, Salge 2010).

Returning to social and cognitive psychology, we find multiple strands of research which also applies a similar logic at the citizens' level. "Spontaneous" attributional activity is defined by Wiener (1985) as an individual's attempt to explain realized outcomes by identifying the cause for them. Individuals try to understand events they are exposed to, and this process is largely an attributional one. For the purpose of our study, we can view performance information about an organization as an outcome to which the individual can generate attributions. Wiener's review of a large number of existing studies shows that negative events elicit much more spontaneous attributional activity than positive events. For instance, Gilovich (1983), involved participants in an experimental bet on sporting events. Subsequently, the participants shared their thoughts about the game, and those who had lost spent more time discussing the game than those who won. A related strand of research focuses on individuals "*counterfactual thinking*", which denotes a mental representation of alternatives to the past (Roese 1997). It finds that counterfactual thinking is more likely to be triggered by negative events and emotions. Roese (1997, 133) stresses that counterfactual thinking does not require deliberate activation but is prompted automatically when we are faced with negative input. We here see a clear analogy to the view of negative performance in organizational theory as trigger for learning. However, here failure is a seen as a fundamental human signal to look for antecedents for the observed outcome. This provides the basis for our second main hypothesis:

Spontaneous attributions from negative performance: *Negative performance information will elicit spontaneous attributions to a greater degree than positive performance information of the same magnitude*.

Design: Three Experimental Studies in Two Nationally-Representative Samples

In order to test our two central hypothesis on the nature of the negativity bias we have conducted three studies embedded in two independent surveys. Study 1 tests the negativity bias on the basis of our modified equivalence framework by comparing the effect on citizens of positive and negative performance information in relation to neutral information containing both positive and negative aspects. Study 2 is a replication of one of the vignette experiments in study 1 in order to confirm the magnitude and validity of the negativity bias. Study 3 investigates differences in how negative, neutral, and positive performance information leads to spontaneous attribution of responsibility, as expected in our second main hypothesis. The three studies are reliant on two large and nationally-representative samples of Danish citizens. Studies 1 and 3 are reliant upon data from the same survey collected in March 2015 (n=1,559) and the replication in study 2 draws upon a survey conducted in May 2015 (n=1,013). Specific aspects of each survey will outlined independently as each study is presented. Both experimental surveys were fielded in YouGov's Danish online panel.[1] Both were restricted to citizens between the ages of 18 and 74 and were pre-stratified on gender, geographical region, age, and political party choice in order to achieve a near-representative sample of the Danish population. Descriptive statistics for both samples are provided in appendix table A1. Population representative samples provide strong external validity for the findings to the public-at-large. Moreover, the large sample sizes (and relative few experimental conditions) mean that we have statistical power to detect even relatively small degrees of negativity bias.

[1]As a payment for participating in the survey, participants received a number of "points," which can be used for purchases in YouGov shop.

Study 1: A Test of the Negativity Bias in Performance Information Evaluation

The purpose of the first experimental study is to estimate a negativity bias with an equivalence framing experiment that includes a neutral category. Here the negativity bias estimate is equal to the difference between the negative vs. neutral treatment and the positive vs. neutral treatment. If the former is significantly larger than the latter, there is evidence of a negativity bias.

Participants

The experiments conducted in study 1 are reliant upon data from the first nationally-representative sample (n=1,559), which took around five minutes to complete and included a number of experiments and background questions.[2] At the onset of the survey all participants were informed about that the study would include a number of hypothetical cases - which nonetheless resemble real-life scenarios and cases.

Experimental Design and Procedure

Study 1 consists of two experimental vignettes focusing on hospital and school performance. Both vignettes are between-subject design with four different treatments in each. The treatments are outlined in table 1 and 2. Both vignettes are equivalence framing experiments where conditions contain logically-equivalent information but differ in the valence of the performance information as either positive, neutral, or negative (Levin, Schneider, and Gaeth 1998; Tversky and Kahneman 1981). Both experiments ask subjects to evaluate an unnamed public organization (hospital and school) after having been given some simple piece of performance information. In the hospital case negative and positive treatment is defined by presenting logically-equivalent levels of patient (dis)satisfaction as either a *satisfied* or

[2]The full survey contained 3,145 observations of which, on a subset of 1,559, were assigned to the experiments reported on here. The complete survey was assigned to a sample of 15,535 which yields a response rate of 20.2%.

dissatisfied percentage (cf. table 1). In the school case positive and negative treatment is defined by presenting logically-equivalent information on students' exam results as being either the *pass* rate or the *fail* rate. The numerical content allows us to hold informational content equally across treatments while varying the informational valence between descriptive labels with negative valence ("dissatisfaction" and "fail rate") and positive valence ("satisfaction" and "pass rate").[3] The treatment conditions, as viewed by the subjects in the survey, are shown in appendix figures 1A and 2A.

Table 1: Hospital Performance Information Experiment

Baseline question (all participants):			
Imagine that a hospital has achieved the following results:			
Treatment frames (randomly assigned):			
Positive (n=519)	Neutral I (n=257)	Neutral II (n=264)	Negative (n=529)
78% of the patients are satisfied with their treatment.	*78% of the patients are satisfied with their treatment.* *22% of the patients are dissatisfied with their treatment.*	*22% of the patients are dissatisfied with their treatment.* *78% of the patients are satisfied with their treatment.*	*22% of the patients are dissatisfied with their treatment.*
How well do you think the hospital is doing given the above?			

[3]The specific percentages and rates are chosen as they are fairly realistic in terms of what could be experienced in the real world.

Table 2: School Performance Information Experiment

Baseline question (all participants):			
Imagine that a public school has achieved the following results:			
Treatment frames (randomly assigned):			
Positive (n=519)	Neutral I (n=257)	Neutral II (n=264)	Negative (n=530)
85% of students have passed their exams.	*85% of students have passed their exams.* *15% of students have failed their exams*	*15% of students have failed their exams.* *85% of students have passed their exams.*	*15% of students have failed their exams.*
How well do you think the school is doing given the above?			

In the neutral treatments, subjects are provided with "both sides of the coin", namely both the satisfied/dissatisfied percentage and the pass/fail rate. By introducing a neutral treatment we can now capture the negativity bias by comparing the difference-in-difference between positive-neutral treatments and neutral-negative treatments. In both vignettes the neutral category is split into two separate treatments: one in which the positive valence information is presented first and one in which the negative valence information is presented first. This is conducted in order to make the neutral treatment robust to order effects in the presentation of information (Hogarth and Einhorn, 1992).[4] The response scale was a 101-point scale from 0 to 100 ranging from "very poor" (0) to "very good" (100) for both experiments.[5] In the hospital case the average evaluation across all treatments was 45.3 (SD=25.1), and in the school experiment 48.6 (SD=26.8).[6] In order to confirm that the randomization of subjects to treatment conditions achieved we conducted a set of successful balance tests on the subjects' background characteristics, as reported in appendix table A2.

[4]We could both suspect that primacy or recency effects could affect subjects and thereby pollute the neutrality of the information. In the analysis, the neutral treatments will be treated as one combined neutral treatment which levels out any order effects by pooling the two neutral treatments.
[5]Subjects were not respond "don't know" but had to provide an evaluation of the organization.
[6]The median response time in the hospital case was 15.6 seconds and 15.1 seconds in the school experiment.

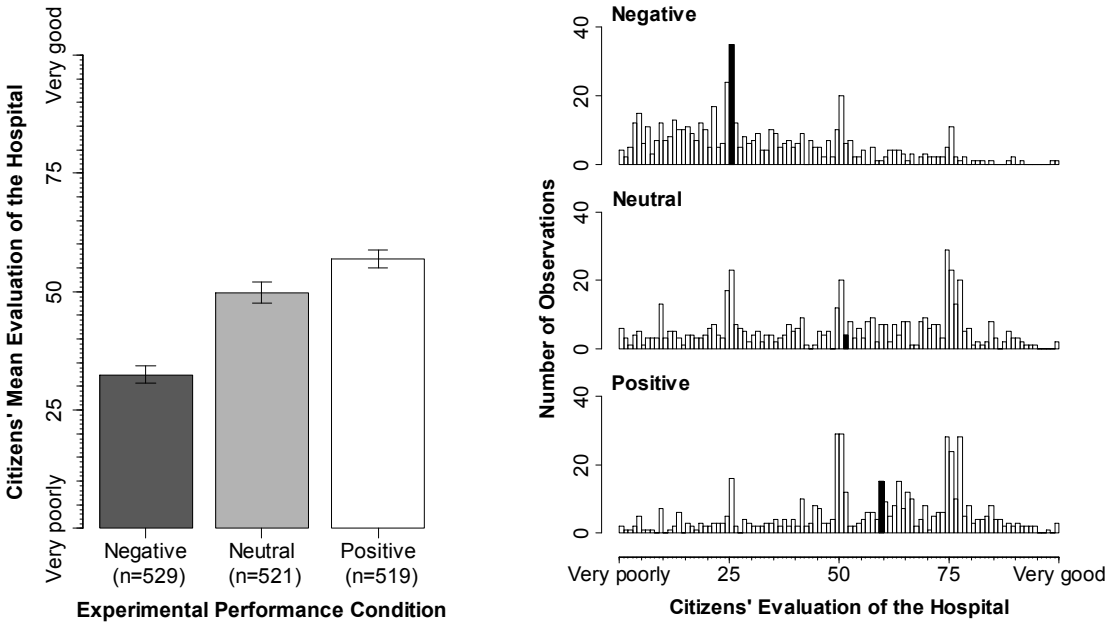
Empirical Findings in Study 1

First we turn to the hospital experiment as shown in figure 1.[7] This reveals major differences in average evaluation and the overall distribution of responses across the three treatment conditions. In the negative performance information treatment, the average evaluation is 32.4 (SD=21.1) which is substantially and significantly lower than both the neutral condition of 49.8 (SD=25.3, $p<0.001$)[8] and the positive treatment of 56.9 (SD=22.2, $p<0.001$). Furthermore, the positive condition is significantly larger than the neutral condition ($p<0.01$). These differences are also mirrored in the distributions, with median values of 26 (negative), 52 (neutral), and 60 (positive). The difference of 24.5 points between the positive and negative condition is whole standard deviation: Providing citizens with substantively identical information but with different valence causes dramatic shifts in their evaluations. More importantly, a substantive part of this effect is driven by a negativity bias. The negativity bias is captured as the difference-in-difference between the negative/neutral condition comparison and the positive/neutral treatment. The results show that the neutral condition was much closer to the positive treatment than the negative one - which is consistent with the negativity bias. Evaluation for the positive frame is 7.1 points higher than the neutral frame while the evaluation for the negative condition is 17.4 points lower than the neutral condition. The valance asymmetry is very clear: If we compare this difference-in-difference directly we find a substantial and significant negativity bias as the negative frame is 10.3 points further away from the neutral condition than the distance between the neutral and the positive ($p<0.001$). Taken together, we find convincing evidence of a negativity bias for the hospital experiment in support of our first hypothesis.

[7]All p-values are computed with t-tests for means or proportions unless some other test is explicitly noted in the text.

[8]As noted in the methods section, we have polled the two neutral conditions which rated the positive and negative information. The first neutral condition had average of 51.7 points (SD=25.0) and the second natural condition was evaluated at 48.0 points (SD=25.5). This indicates that listing the positive valence information before the negative valence information induced slightly higher evaluations. However, overall the difference between the neutral conditions is very small and insignificant ($p=0.10$).

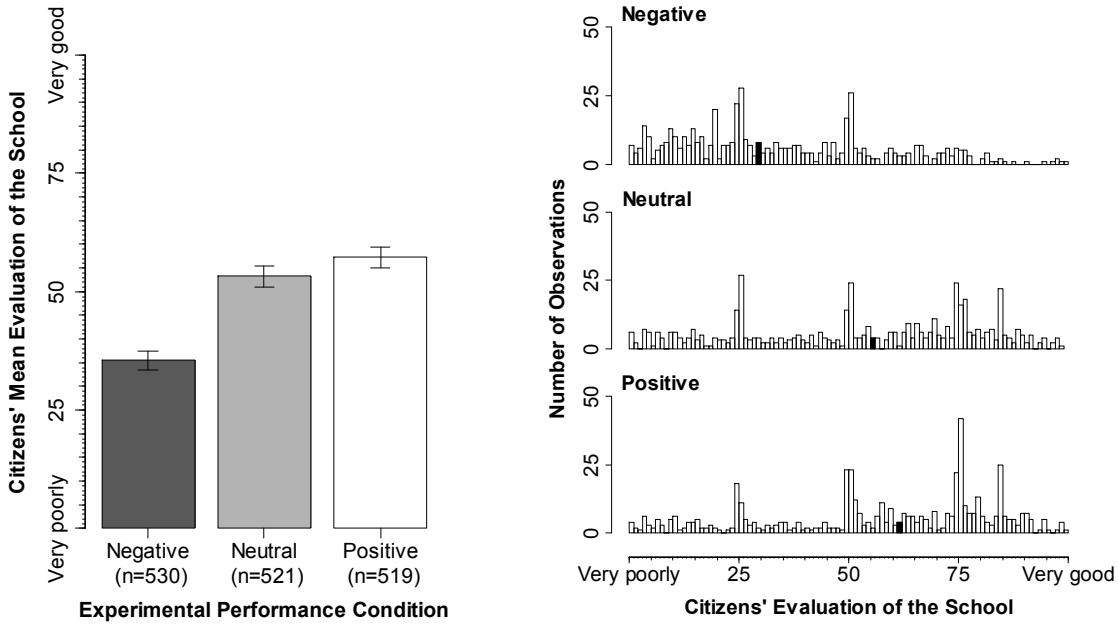
Figure 1: Effect of Randomized Performance Information from a Hospital on Citizens' Evaluation of Public Service Performance (N=1559). *Note: The left panel shows the mean evaluations across the three experimental conditions with 95% confidence intervals. The right panel illustrates the distribution of evaluations across the three conditions with the median evaluation in black.*



Following this, we can compare the effects in the school experiment in figure 2. Here the difference appears even larger than in the school case, with the negative condition of a 15% fail rate dramatically different from the two others. The negative performance information treatment with an average of 35.5 (SD=23.0) is substantially and significantly lower than both the neutral treatment of 53.3 (SD=26.6, $p<0.001$)[9] and the positive treatment of 57.3 (SD=25.6, $p<0.001$). Evaluations for the positive treatment are significantly higher than the neutral ($p<0.05$). These differences are also reflected in the median values: 30 (negative), 56 (neutral), and 62 (positive). Again we capture the negativity bias as the difference-in-difference between the negative/neutral condition comparison and the positive/neutral comparison.

[9]As in the hospital experiment the two neutral conditions are combined. The first neutral condition had average of 54.6 points (SD=26.5) and the second natural condition was evaluated at 52.0 points (SD=26.7). Again a slightly higher evaluation for the neutral condition that presents positive valence information first, but, as in the hospital experiment, the difference is not significant (p=0.25).

Figure 2: Effect of Randomized Performance Information Valence from a School on Citizens Evaluation of Public Service Performance (N=1,559). *Note: The left panel shows the mean evaluations across the three experimental conditions with 95% confidence intervals. The right panel illustrates the distribution of evaluations across the three conditions with the median evaluation in black.*



Just from comparing the relative distances between the three conditions we can sense that the negativity bias is even stronger in the school experiment. The difference between the negative and neutral condition was 17.8 points while the difference between the positive and neutral was just 4.0 points. The difference-in-difference is therefore large and significant at 13.8 points ($p<0.001$): how much further away the negative condition is from the neutral. Again very robust evidence of an asymmetry in citizens' evaluations which conforms to a very strong and pervasive negativity bias. Taken together, findings across both experiments are consistent with the expectation of valence-consistent shifts in citizens' evaluations as the valence of the performance labels changes, while the underlying information is held constant. Most importantly, we find that convincing evidence of a negativity bias: there is a persistent asymmetry in the evaluations of positive and negative frames in relation to the neutral one. Evaluations of the negative frame a much further below evaluations for the neutral frame than evaluations of the positive frame are above the neutral frame.

Study 2: Replication of Study 1

Study 1 provided surprisingly strong evidence of a negativity bias in citizens' evaluation of public services from performance information. In light of the more mixed findings on a negativity bias in the existing literature on performance information, we conducted a direct replication study in order to confirm the magnitude of the findings in study 1. We only replicated the school experiment as it showed the largest effect in study 1.

Participants

As in study 1, the replication relies on a large nationally-representative sample of the Danish population (n=1,013). The data was collected between May 21st and the 26th in 2015. A total of 3,502 subjects were invited to participate in the survey yielding a response rate of 28.9% for full completes.[10] The survey contained multiple unrelated experiments with the replication study placed at the very end of the questionnaire. The median response time for the full survey was around 5 minutes.

Experimental Design and Procedure

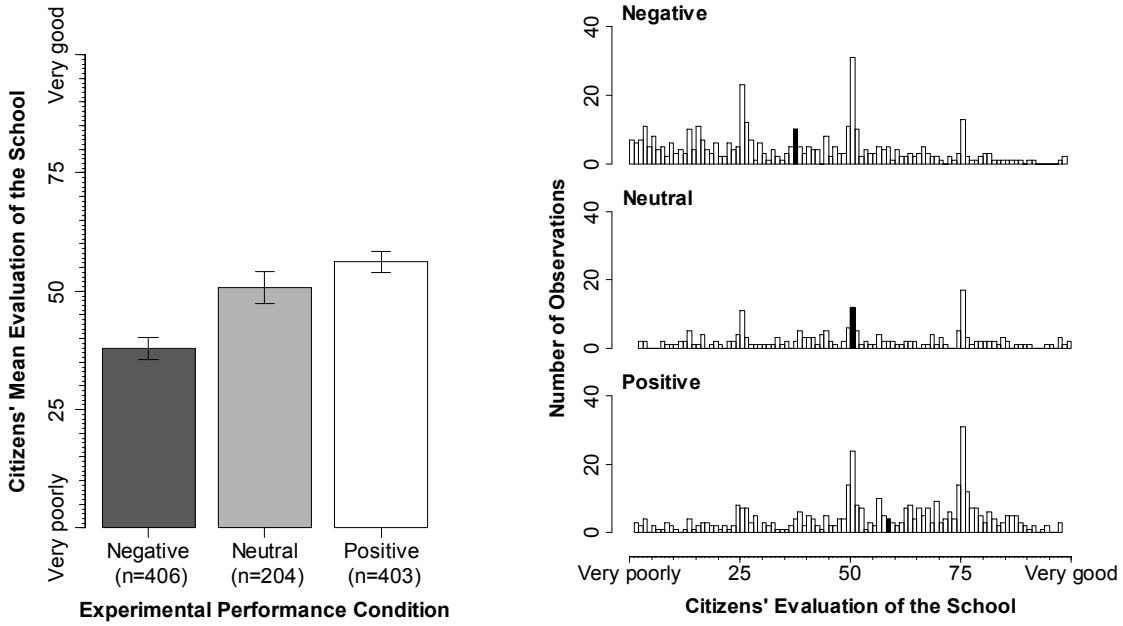
Study 2 is an exact replication of the school experiment in study 1 and relied on the exact same treatment conditions for the school experiment as shown earlier in table 2.[11] As the sample was slightly smaller the number of subjects across the four conditions was: n=406 (negative), n=204 (neutral I and II), and n=403 (negative). The response scale was also exactly the same and the mean response across all conditions was 47.8 (SD=25.1).

[10] Participating subjects received 25 points eligibale for purchases or lotteries at YouGov.
[11] The presentation of the conditions was also the exact and is shown in appendix B.

Empirical Findings in Study 2

Results from the replicated school experiment are shown in figure 3. The overall pattern closely resembles the findings in study 1. The negative condition generates by far the lowest evaluation at 37.8 points (SD=23.7) which is significantly lower than the 50.8 points (SD=24.7) for the neutral frame ($p<0.001$)[12] and the 56.2 points (SD=23.1) for the positive frame ($p<0.001$). Evaluations for the neutral frame are also significantly lower than for the positive frame ($p<0.01$). Median evaluations are 37.5 points for the negative, 51.0 points for the neutral, and 59.0 the positive. Again, we therefore find a strong negativity bias, with the difference between the negative and neutral condition amounting to 13 points - while the positive condition is just 5.4 points below the neutral condition. This provides a difference-in-difference between conditions of 7.6 points ($p<0.05$) which indicates a credible negativity bias. Taken together, study 2 replicates the findings of study 1.

Figure 3: Effect of Randomized Performance Information Valence about a School on Citizens Evaluations (n=1,013). *Note: The left panel shows the mean evaluations across the three experimental conditions with 95% confidence intervals. The right panel illustrates the distribution of evaluations across the three conditions with the median evaluation in black.*



[12]As in the original study there is no substantive or significant difference between the two polled natural conditions (p=0.36).

Study 3: Performance Valence and Attribution of Responsibility

In study 3 we proceed to test differences in spontaneous attributions of responsibility if exposed to negative, neutral, and positive performance information. As argued in our second hypothesis, we expect that negative performance information to induce spontaneous attributions about responsibility among citizens (Weiner 1985; Roese 1997).

Participants

The participants are the same as those who participated in study 1 (n=1,559).

Experimental Design and Procedure

We were reliant on a question that immediately followed the treatment conditions in Study 1. Here the subjects were confronted with open-ended questions and a text box on a separate page. Subjects were asked to elaborate on their thoughts about the hospital and school which they just evaluated. Specifically, subjects were asked to: *"Write a comment on what you immediately think about the hospital given the provided information:"* Following the school experiment the subjects were asked the exact same question with regards to the school.[13] We relied on open-ended responses because we wanted to capture spontaneous attributions of responsibility. In closed-ended questions the response scale would provide strong cues to subjects about the intent of the survey (RePass 1971; Geer 1991). This would have been problematic as we did not want to create demand effects among subjects which provided hints as to attribution of responsibility. For instance, it is important but at the same time difficult to avoid words like "blame" or "credit" when measuring responsibility attributions (Rudolf 2003). Opting for open-responses reduces demand effects and allows us to study the extent to which attributions arise spontaneously. In order to analyze the responses we relied

[13]The screen, as viewed by each subject, is provided in appendix figure A3. Subjects had to fill in at least one character in order to proceed in the survey. If they attempted to proceed without providing a response then they were informed about this response rule in a brief notification on the screen.

in on manual coding of the responses. The coder had no prior knowledge of the experiments and was only informed that the text comments were responses from citizens after exposure to performance information. In the coding process the coder was blinded to the treatment conditions and was presented with the open-text responses in random order. The primary variable to be coded is an indicator of attribution of responsibility. Specifically, the coding manual instructed the coder to answer the following question for each entry into the open text box: "*Does the text provide one or more considerations about possible explanations or reasons for the school's/hospital's accomplished result?*" This provided us with a dummy variable indicating text responses with spontaneous attributions. We here relied on the standard coding in psychological research on attributions in written material, defined as, "*a phrase or sentence in which some performance outcome…was linked with a reason for that outcome*" (Bettman and Weitz 1983, 173).[14] Descriptive statistics of the text responses and coding is reported in table 3. Subjects spent around 30 seconds providing their text response and produced an average of about 12 words for each experiment. Most importantly, about 40% of the subjects' responses were coded as containing attributional content. In the following analysis, our main goal is to understand variation in the attributional activity across the three experimental groups. Following this, we will also provide some examples of the attributions made by subjects. Non-attribution content ("other responses") usually took the form of elaborating on the closed responses which subjects provided in study 1. Finally, the valence of attributions was also coded as either negative, neutral, or positive. Given our hypothesis we are mostly interested in negative valence attributions which are likely to contain blame content. On average between 60% and 70% of the attributions in the two experiments were of negative valence.

[14]This definition is fairly typical throughout existing studies on spontaneous attributions (Weiner 1985, 76).

Table 3: Descriptive Statistics, Coding, and Reliability in Study 3

	Hospital experiment	School experiment
Descriptives:		
Mean number of words in responses	11.9	12.1
Median response latency (seconds)	32.3	32.4
	n=1,559	*n=1,559*
Attribution coding:		
Percentage of responses with attributions	41.1%	39.3%
Percentage of other responses	52.6%	55.3%
Percentage with no response	6.3%	5.4%
Total	100%	100%
	n=1,559	*n=1,559*
Reliability of attribution coding:		
Percent of inter-coder agreement	91.7%	82.8%
Cohen's Kappa	0.82	0.63
	n=157	*n=157*
Coding of attribution valence:		
Percentage of attributions with negative valence	60.3%	70.1%
	n=635	*n=604*

Another independent coder coded a random sample of 10% of the observations (n=157) for both the hospital and school experiments in order to test for an inter-coder reliability test. In both experiments the level of mutual agreement is above 80%. Cohen's Kappa captures inter-coder agreement for qualitative items and accounts for agreement attributable to chance (Blackman and Koval 2000). We find high to very high Kappa values for the coding of attributions in both experiments.[15]

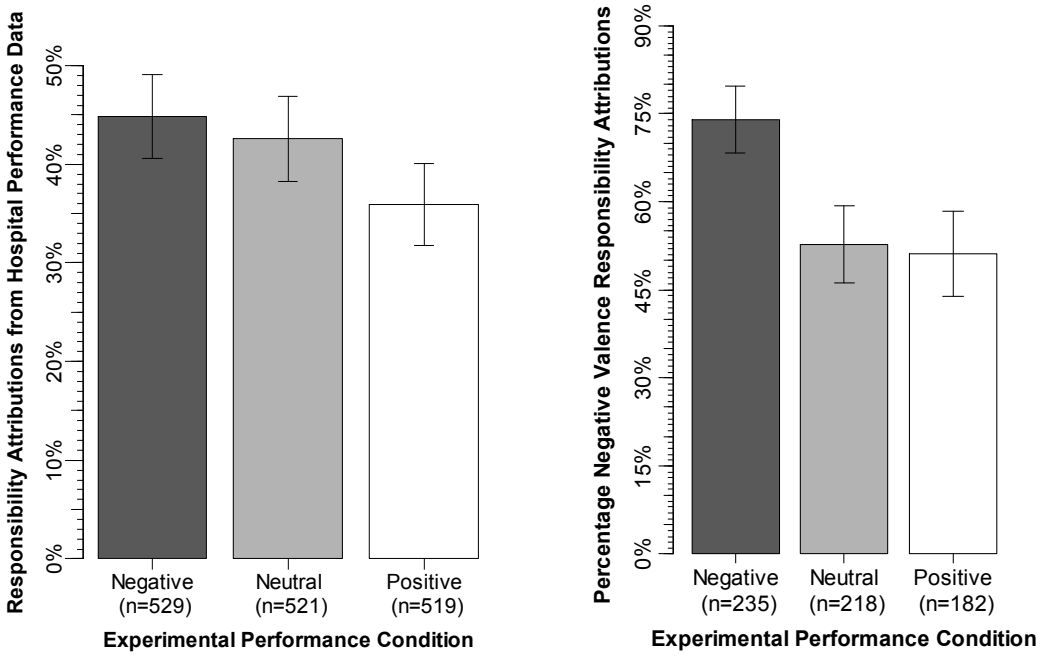
Empirical Findings in Study 3

The main findings from the hospital experiment are presented in figure 4 as the percentage of open-ended responses containing attributions of responsibility across the three conditions. We find the highest degree of attribution in the negative condition at 44.8% which is not significantly higher than the 42.6% for the neutral condition (p=0.39), but substantively and

[15]Kappa values of 0.6 or higher are regarded as a substantial level of mutual agreement (Landis and Koch 1977).

significantly higher than the just 35.9% of subjects engaging in attribution of responsibility in the positive condition ($p<0.01$).[16] The difference between the positive and negative condition of around 9 percentage points amounts to a difference in attributions of about 25% which suggests that negative performance is indeed a trigger for spontaneous attribution of responsibility relative to positive performance. Moreover, the right panel shows that 74% of attributions under the negative treatment condition are of negative valence, which is significantly more than for 52.8% in the neutral condition ($p<0.001$) and 51.1% in the positive condition ($p<0.001$).[17]

Figure 4: Effect of Randomized Performance Information about a Hospital on Citizens Attribution of Responsibility. *Note: The left panel shows the percentage of open-ended text responses containing attributions about performance across the three experimental conditions (with 95% CIs). The right panel shows the percentage of negative valence attributions of the total number of attributions across the three experimental conditions (with 95% CIs).*



[16]The degree of responsibility attribution also significantly lower in the positive condition than in the neutral condition ($p<0.05$).

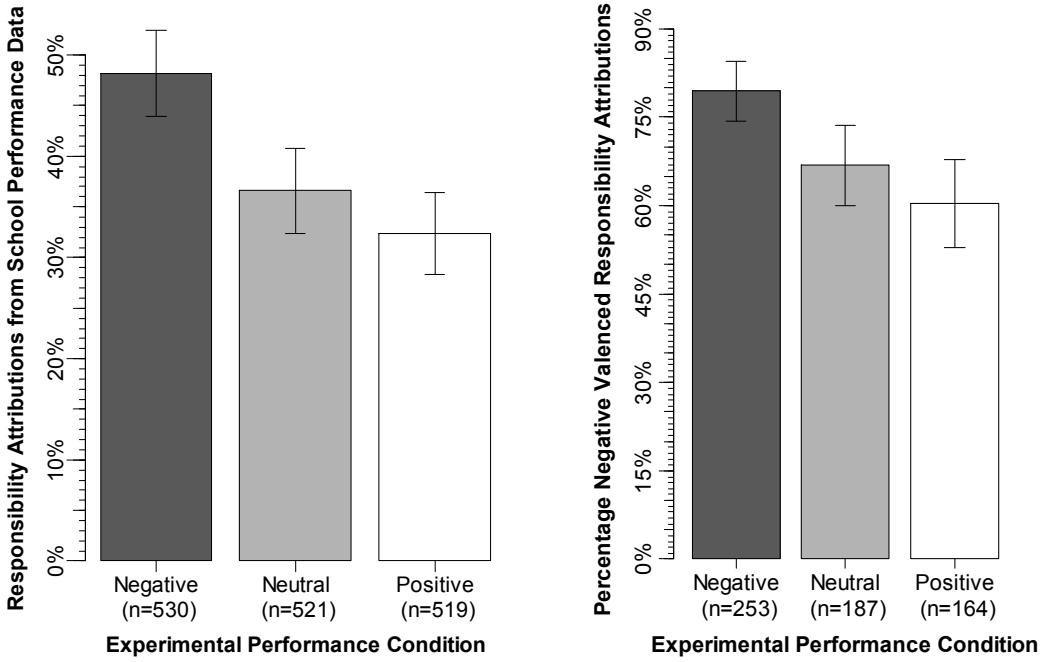
[17]The percentage of negative valence attribution is not significantly different from each other in the neutral and positive conditions ($p=0.74$).

If we take an exploratory look at some examples of attributions made be the subject, we find that most of them are directed at hospital level actors like the management, doctors, or nurses. One example states how "bad doctors, and nurses that care more about coffee meetings than doing their job." Some subjects provide quite detailed attributions: "If 1 out of 5 patients are unsatisfied, one needs to determine what they are complaining about. What is the commitment of the staff? Where is the management in this particular case?" Others are less subtle in their attributions of responsibility: "there has to be a bad management at that hospital."

Findings from the school experiment are presented in figure 5. Again, the negative information condition produces the greatest degree of responsibility attribution by far, with 48.2% of responses. It is significantly higher than for both the neutral condition (36.6%, $p<0.001$) and the positive condition (32.3%, $p<0.001$). The neutral and positive conditions are not significantly different from each other ($p=0.15$). The difference between the positive and negative conditions of about 16% reveals a difference in spontaneous attributions of about 49%. Again we see in the right panel that attributions are predominantly negative - in particular under the negative condition. In the negative condition 79.4% of the attributions are negative, which is significantly higher than for the 66.8% for the neutral condition ($p<0.01$) and the 60.4% ($p<0.001$).[18] In the school experiment we also see that most attributions are directed towards the school level. Many of them are directed at the teachers: "the commitment among the personnel might be too small." Others in this category point to failure of the management, i.e. "poorly managed school." Unlike the hospital experiment, we find examples of attributions directed towards the users of the service such as students and parents: "The parents refrain from taking responsibility. It is much easier just to blame the school system, but there will always be some that cannot make it."

[18] The percentage of negative valence attribution is not significantly different from each other in the neutral and positive conditions ($p=0.21$).

Figure 5: Effect of Randomized Performance Information about a School on Citizens' Attribution of Responsibility. *Note: The left panel illustrates the percentage of open-ended text responses containing attributions about performance across the three experimental conditions (with 95% CIs). The right panel shows the percentage of negative valence attributions of the total number of attributions across the three experimental conditions (with 95% CIs).*



Conclusions and Implications

This article grew out of the recent interest in how the negativity bias affects how citizens draws inferences from performance information. We found that the negativity bias has a strong and multifaceted impact on what inferences citizens draw from performance information regarding public services. First, we found a very robust experimental negativity bias in citizens' evaluations of negative, neutral, and positive performance information. Negative performance information induced very negative evaluations of two very different public services, while neutral and positive performance information led to a considerably larger amount of similar evaluations. The asymmetrical effects were large and the equivalence framing experiment provided a conservative test which, together with the independent replication, gave strong support for the negativity bias. Second, we provided the

first experimental evidence of spontaneous attribution in citizens' response to negative performances. After being exposed to negative valence performance information citizens were much more likely to spontaneously produce causal attributions about performance. Furthermore, the attributions made following negative performance information were much more likely to contain negative content in the form of "blame" for the outcome.

It is important to stress the limits of the studies and findings presented here. While the external validity across subjects is strengthened by the use of two nationally-representative samples, we should always be cautious when extending results from vignette experiments embedded in a survey to the messy reality of how citizens are exposed to performance information in the real world. A key focus for follow-up studies should be to test the external validity of the negativity bias in real-world administrative settings.

The findings have a number of implications for both public managers and future research. First, they provide a robust piece in the complex puzzle of understanding the inferences that citizens draw from performance information. Identifying the negativity bias is interesting it its own right because it corrects our view of how citizens respond to performance information. The average citizen's response to performance information is largely driven by the excessive weight given to negative information which not only affects evaluations but also attributions about performance. Some have speculated if the negativity bias is a "constant in human affairs" or if it only emerges in public administration due to certain institutional features which amplify negative information (Hood 2007; 2010, 13). Our experiments point to the fact that the bias is a very fundamental feature of human information processing which naturally affects how performance information is perceived. This raises the question of how the negativity bias works if it is in conflict with other fundamental determinants of how citizens process performance information. For instance, a natural extension would be to integrate the negativity bias with recent findings on how motivated

reasoning and implicit attitudes guide the inferences which citizens draw from performance data (Marvel 2014; Bækgaard and Serritzlew 2015). This would allow us to study how citizens deal with negative performance information if they also are motivated by pre-existing attitudes to have a favorable view of the organization.

Secondly, the findings offer credibility to theories on blame avoidance which posits the negativity bias among citizens as being the underlying mechanism as to why politicians pursue a minimax strategy (Weaver 1986, Dixon and Hood 2010, Moynihan 2012). Moreover, the fact that negativity not only drives evaluations of performance information but also triggers attributions about responsibility draws an even more powerful link between the negativity bias and blame avoidance as a downstream effect. Managers and politicians have at least two reasons to be concerned with negative performance information: (a) It has a stronger effect on evaluations in negative direction and (b) it is more likely to elicit attributions which negatively link poor performance with those responsible. At the same time, the results call for greater inquiry into how politicians and managers avoid blame – in particular how performance information can be turned to from a mechanism for appointing blame into one of blame avoidance. As Hood (2010, 48) forcefully argues: "victims of negativity bias" will go to great lengths to **"**correct such bias to keep blame at bay". Here the findings direct our attention to how the effect of negative performance information may be mitigated with "presentational strategies" (Hood 2010, 17) or attempts to "redefine the issue" (Weaver 1986, 386). The equivalence framing experiment illustrates how, with subtle changes in labeling, the very same pieces of performance information can cause major shifts in citizens' evaluations and attributions. It calls upon us to look closer at observational data of how organizations present information with an aim of shaping the public's perception of performance .

Finally, we need more fundamental research regarding what kind of "bias" a negativity bias would be (Soroka 2006; 2014). The "bias" in negativity bias can be understood in at least two ways: First, in a minimal sense we can understand bias as a systematic divergence from a model of the world which we rely on. In this sense, the notion of a bias helps correct the false (layman's) impression that positive and negative information is given symmetrical or equal weight. A second interpretation of "bias" would make a value-laden statement. It would emphasize that an over-reliance on negative input leads to a less "optimal" judgment compared with some normative models of rational decision-making which tell us to put equal weight to positive and negative information of the same magnitude. The findings support the first interpretation, but the second interpretation would require direct comparisons with alternative ways that citizens could understand performance data. In public administration we often implicitly draw on the latter interpretation which leads to a concern about how to make citizens value "good" performance. We believe that correcting the negativity bias could help turn organizations' focus away from avoiding failure and towards the pursuit of excellence. This may very well be true, but from a model of bounded rationality, we must also keep the possibility open that the negativity bias could be a rule of thumb that, in most instances, serves citizens well. In short, it might place an over-reliance on failure detection, but in the aggregate provides citizens with a fast and frugal decision-rule for making inferences about public sector performance. Moreover, our finding of increased attributional activity for negative information offers a parallel to those who stress the functional basis of counterfactual thinking in psychology (Roese 1997). It implies that citizens think more about causes and effects in the realm of public sector performance, which we mostly would recognize as a good and productive thing. To summarize, the negativity bias is real and exerts a strong influence on how citizens draw on performance information. We have yet to find out if this a good or a bad thing.

References

Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. 2001. Bad is Stronger than Good. *Review of General Psychology* 5(4): 323–370.

Bettman, James R., & Weitz, Barton A. 1983. Attributions in the board room: Causal reasoning in corporate annual reports. *Administrative science quarterly* 28: 165–183.

Blackman, Nicole, and John Koval. 2000. Interval Estimation for Cohen's Kappa as a Measure of Agreement. *Statistics in Medicine* 19(5): 723–41.

Boyne, George A., Oliver James, Peter John, and Nicolai Petrovsky. 2009. Democracy and Government Performance: Holding Incumbents Accountable in English Local Governments. *Journal of Politics* 71(4): 1273–1284.

Bækgaard, Martin and Søren Serritzlew. 2015. Interpreting performance information: Motivated reasoning or unbiased comprehension. *Public Administration Review*.

Charbonneau, E., & Bellavance, F. (2012). Blame avoidance in public reporting: Evidence from a provincially mandated municipal performance measurement regime. *Public Performance & Management Review* 35(3), 399–421.

Cyert, Richard M., and James G. March. 1963. *A behavioral theory of the firm.* Englewood Clis, New Jersey: Prentice-Hall International Series in Management and Behavioral Sciences in Business Series.

Dixon, R., Arndt, C., Mullers, M., Vakkuri, J., Engblom-Pelkkala, Kristiina & Hood, Christopher. 2013. A lever for improvement or a magnet for blame? Press and political responses to international educational rankings in four EU countries. *Public Administration* 91(2), 484–505.

Fiske, Susan T. 1980. Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology* 38(6): 889–906.

Geer, John G. 1991. Do open-ended questions measure "salient" issues? *Public Opinion*

Quarterly 55(3), 360–370.

Gilovich, Thomas. 1983. Biased evaluation and persistence in gambling. *Journal of personality and social psychology* 44(6): 1110–1126

Hibbing, John R., Kevin B. Smith, John R. Alford. 2014. Differences in Negativity Bias underlie Variations in Political Ideology. *Behavioral and Brain Science* 37: 297–350.

Hogarth, R. M., & Einhorn, H. J. 1992. Order effects in belief updating: The belief-adjustment model. *Cognitive psychology* 24(1): 1–55.

Hood, Christopher. 2007. What Happens When Transparency Meets Blame-avoidance? *Public Management Review* 9(2): 191–210.

Hood, Christopher. (2010). *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton University Press.

Hood, Christopher, and Ruth Dixon. 2010. The political payoff from performance target systems: no-brainer or no-gainer? *Journal of Public Administration Research and Theory* 20(2): 281–298.

James, Oliver, and Alice Moseley. 2014. Does Performance Information about Public Services Affect Citizens' Perceptions, Satisfaction, and Voice Behaviour? Field Experiments with Absolute and Relative Performance Information. *Public Administration* 92(2): 493–511.

James, Oliver. 2011a. Managing Citizens' Expectations of Public Service Performance: Evidence from Observation and Experimentation in Local Government. *Public Administration* 89(4): 1419–1435.

James, Oliver. 2011b. Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments. *Journal of Public Administration Research and Theory* 21(3): 399–418.

James, Oliver, and Peter John. 2007. Public Management at the Ballot Box: Performance

Information and Electoral Support for Incumbent English Local Governments. *Journal of Public Administration Research and Theory* 17(4): 567–580.

Johnsen, Åge. 2012. Why Does Poor Performance Get So Much Attention in Public Policy? *Financial Accountability & Management* 28(2): 121–142.

James, Oliver, & John, Peter. 2007. Public management at the ballot box: Performance information and electoral support for incumbent English local governments. *Journal of Public Administration Research and Theory* 17(4): 567–580.

Kelley, Harold H. 1973. The processes of causal attribution. *American psychologist* 28(2): 107–128.

Landis, J. Richard, and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1): 159–74.

Levin, Irwin P., Sandra L. Schneider, and Gary J. Gaeth. 1998. All Frames are not Created Equal: A Typology and Critical Analysis of Framing Effects. *Organizational behavior and human decision processes* 76(2): 149–188.

Levin, Irwin P., and Gary J. Gaeth. 1988. How Consumers are Affected by the Framing of Attribute Information Before and After Consuming the Product. *Journal of Consumer Research* 16: 374–378.

March, James G., and Herbert A. Simon. 1958. *Organizations*. New York: Wiley

Marvel, John D. 2014. The Boston Marathon Bombings: Who's to Blame and Why It Matters for Public Administration. *Public Administration Review* 74(6): 713–725.

Meier, Kenneth J., Nathan Favero, and Ling Zhu. 2015. Performance Gaps and Managerial Decisions: A Bayesian Decision Theory of Managerial Action. *Journal of Public Administration Research and Theory*.

Moynihan, Donald P. 2012. Extra-Network Organizational Reputation and Blame Avoidance in Networks: The Hurricane Katrina Example. *Governance*, 25(4), 567–588.

Nielsen, Poul. A., & Baekgaard, Martin 2015. Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*, 25 (2): 545–569

Olsen, Asmus L. 2013. Leftmost-digit-bias in an Enumerated Public Sector? An Experiment on Citizens Judgment of Performance Information. *Judgment and Decision Making* 8(3): 365–371.

Olsen, Asmus L. 2015. Citizen (Dis)satisfaction: An Equivalence Framing Study. *Public Administration Review* 75(3): 469–478.

Quattrone, George A., and Amos Tversky. 1988. Contrasting Rational and Psychological Analyses of Political Choice. *American Political Science Review* 82(3): 719–736.

Roese, Neal. J. 1997. Counterfactual thinking. *Psychological bulletin*, 121(1), 133–148.

Rose, Richard. 1991. What is Lesson-Drawing? *Journal of Public Policy* 11(1): 3–30.

Rozin, Paul, and Edward B. Royzman. 2001. Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review* 5(4): 296–320.

Rudolph, Thomas J. 2003. Who's responsible for the economy? The formation and consequences of responsibility attributions. *American Journal of Political Science* 47(4): 698–713.

Salge, Thorsten O. 2011. A behavioral model of innovative search: Evidence from public hospital services. *Journal of Public Administration Research and Theory*,*21*(1), 181–210.

Soroka, Stuart N. 2006. Good News and Bad News: Asymmetric Responses to Economic Information. *Journal of Politics* 68(2): 372–385.

Soroka, Stuart N. 2014. *Negativity in Democratic Politics: Causes and Consequences*. Cambridge University Press.

RePass, David E. 1971. Issue salience and party choice. *American Political Science Review* 65(2): 389–400.

Tversky, Amos, and Daniel Kahneman. 1981. The Framing of Decisions and the Psychology of Choice. *Science* 211(4481): 453–458.

Yang, Kaifeng, & Holzer, Marc. 2006. The performance–trust link: Implications for performance measurement. *Public Administration Review* 66(1): 114–126.

Weaver, R. Kent. 1986. The Politics of Blame Avoidance. *Journal of Public Policy* 6(4): 371–398.

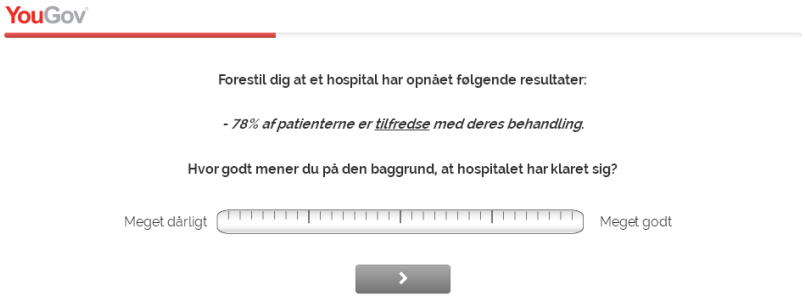
Weiner, Bernard. 1985. "Spontaneous" causal thinking. *Psychological Bulletin* 97(1): 74–84.

APPENDIX

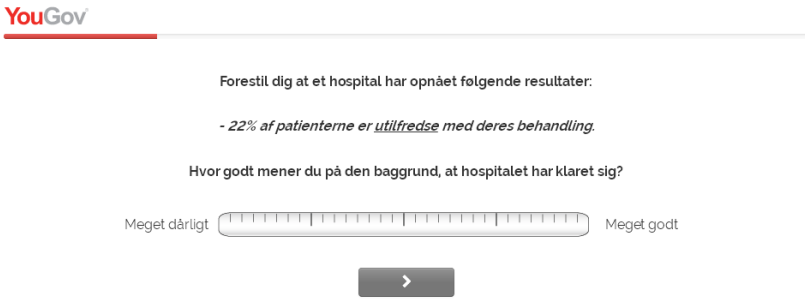
Table A1: Socio-demographic Characteristics of Study Participants

Variable	Study 1 & 3 (n=1,559)	Study 2 (replication) (n=1,013)
Males (%)	48.6	50
Age (years)	47.7	46.6
Geographic region (%)		
Capital area	33.4	32.0
Zealand	16.1	13.8
Southern Denmark	24.2	21.5
Middle Jutland	15.6	22.2
Northern Jutland	10.8	10.5
Education (%)		
High school or lower	21.4	21.2
Short-cycle tertiary	19.3	19.0
Vocational training	12.6	11.5
BA or equivalent	31.0	30.9
MA or higher degree	15.7	17.4
Gov. party supporter (%)	36.1	33.4

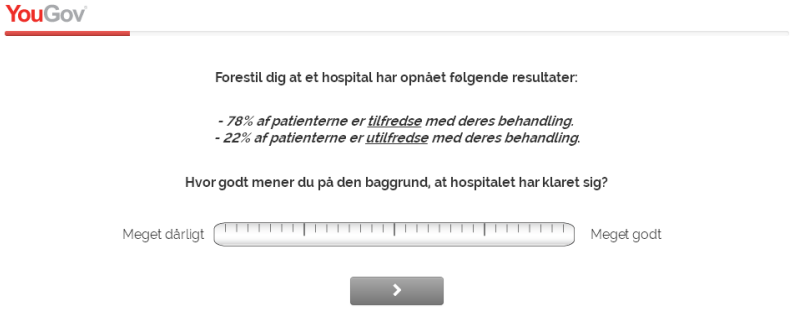
Figure A1: The Hospital Experiment - The Treatment Conditions as Viewed by the Subjects in Study 1.



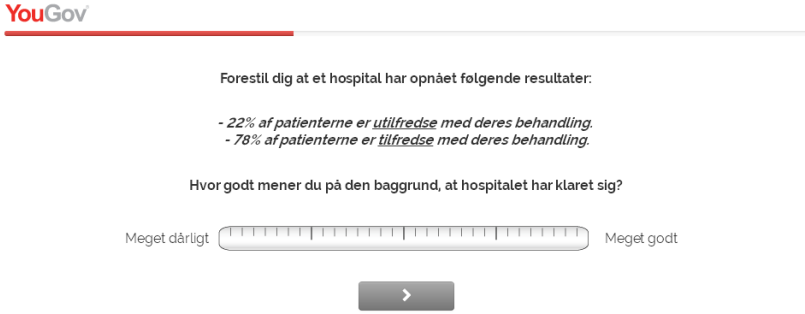
Positive condition: "*Imagine that a hospital has achieved the following results: 78% of the patients are satisfied with their treatment. How well do you think the hospital is doing given the above?"*



Negative condition: "*Imagine that a hospital has achieved the following results: 22% of the patients are dissatisfied with their treatment. How well do you think the hospital is doing given the above?"*



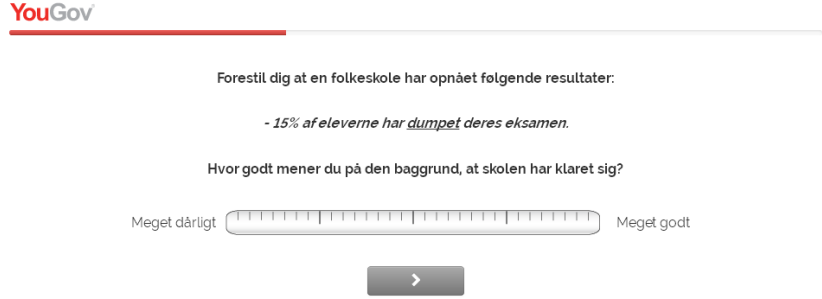
Neutral condition I: "*Imagine that a hospital has achieved the following results: 78% of the patients are satisfied with their treatment. 22% of the patients are dissatisfied with their treatment. How well do you think the hospital is doing given the above?"*



Neutral condition II: "*Imagine that a hospital has achieved the following results: 22% of the patients are dissatisfied with their treatment.78% of the patients are satisfied with their treatment. How well do you think the hospital is doing given the above?"*

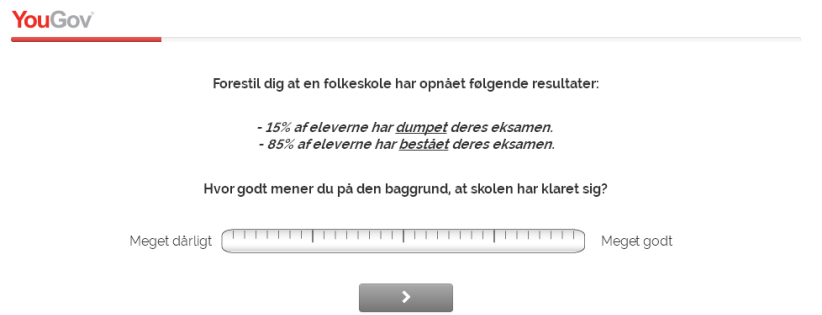
Figure A2: School Experiment: The Treatment Conditions as Viewed by the Subjects in Study 1 and 2 (replication)

Positive condition: "*Imagine that a public school has achieved the following results: 85% of students have passed their exams. How well do you think the school is doing given the above?"*



Negative condition: "*Imagine that a public school has achieved the following results: 15% of students have failed their exams. How well do you think the school is doing given the above?"*

Neutral condition I: "*Imagine that a public school has achieved the following results: 85% of students have passed their exams. 15% of students have failed their exams. How well do you think the school is doing given the above?"*



Neutral condition II: "*Imagine that a public school has achieved the following results: 15% of students have failed their exams. 85% of students have passed their exams. How well do you think the school is doing given the above?"*

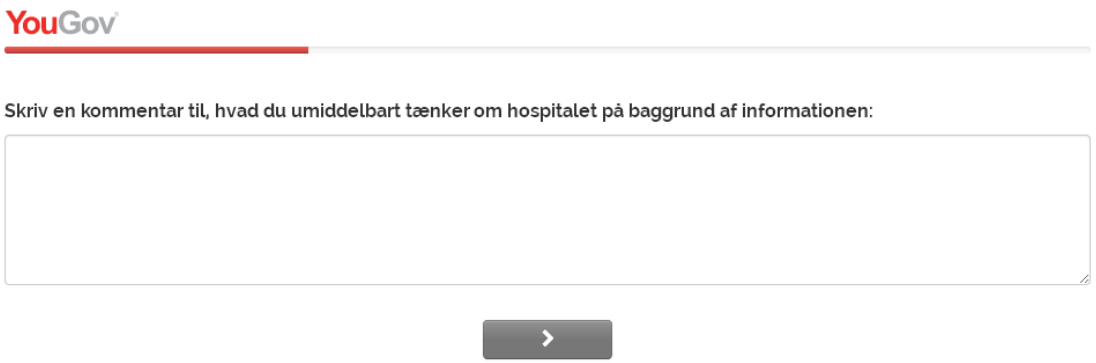
Table A2: Test of Randomization of Treatment Conditions in Study 1 and 3 (n=1,559)

Variable	Negative	Neutral I & II	Positive	*F*	*Prob. > F*
Males (%)	47.4	47.8	50.5	0.58	0.56
Age (years)	47.0	48.2	48.0	1.02	0.36
Geographic region (%)					
Capital area	34.6	32.4	33.1	0.26	0.75
Zealand	16.6	15.9	15.6	0.11	0.90
Southern Denmark	21.6	26.7	24.3	1.89	0.15
Middle Jutland	16.1	14.2	16.6	0.61	0.54
Northern Jutland	11.2	10.7	10.4	0.08	0.93
Education (%)					
High school or lower	21.2	20.0	22.9	0.69	0.50
Short-cycle tertiary	18.9	19.4	19.7	0.05	0.95
Vocational training	12.5	14.0	11.4	0.83	0.44
BA or equivalent	29.3	33.0	30.8	0.85	0.43
MA or higher degree	18.1	13.6	15.2	2.01	0.12
Gov. party supporter (%)	36.1	35.7	36.4	0.03	0.97

Table A3: Test of Randomization of Treatment Conditions in Study 2 – replication (n=1,013)

Variable	Negative	Neutral I & II	Positive	*F*	*Prob. > F*
Males (%)	46.6	51.0	52.9	1.66	0.19
Age (years)	47.8	43.0	47.2	7.19	0.0008
Geographic region (%)					
Capital area	33.7	27.0	32.8	1.53	0.22
Zealand	13.3	17.6	12.4	1.64	0.20
Southern Denmark	20.0	22.1	22.8	0.52	0.60
Middle Jutland	23.9	20.1	21.6	0.64	0.53
Northern Jutland	9.1	13.2	10.4	1.23	0.29
Education (%)					
High school or lower	18.7	27.5	20.6	3.19	0.04
Short-cycle tertiary	22.2	14.6	17.9	2.73	0.07
Vocational training	11.6	12.7	10.9	0.22	0.80
BA or equivalent	29.6	29.4	33.0	0.69	0.50

Figure 3A: The Survey Questions as Viewed by the Subjects in Study 3.



Following the four hospital conditions: "*Write a comment on what you immediately think about the hospital given the provided information:"*

Following the four school conditions: "*Write a comment on what you immediately think about the school given the provided information:"*